



Gryb, S., & Thebault, K. P. Y. (2016). Schrödinger Evolution for the Universe: Reparametrization. *Classical and Quantum Gravity*, 33(6), [065004]. <https://doi.org/10.1088/0264-9381/33/6/065004>

Peer reviewed version

Link to published version (if available):
[10.1088/0264-9381/33/6/065004](https://doi.org/10.1088/0264-9381/33/6/065004)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Schrödinger Evolution for the Universe: Reparametrization

Sean Gryb¹ and Karim Thébault²

¹*Radboud University Nijmegen, Institute for Mathematics,*

Astrophysics and Particle Physics, The Netherlands

²*Department of Philosophy, University of Bristol, Bristol, UK*

Abstract

We develop a new framework for constructing the quantum observables of theories invariant under global time reparametrizations. Our starting point is a generalised Hamilton–Jacobi formalism for totally constrained systems. Our end result is a quantum formalism which simultaneously preserves: i) genuine evolution; and ii) the constraint structure of the classical formalism in the semi-classical limit. We thus offer a novel resolution of the ‘problem of time’ for theories invariant under global time reparametrization. There is the potential for the application of our proposal to the quantization of gravity when understood in terms of the Shape Dynamics formalism.

CONTENTS

I. Introduction	2
II. Symmetry and Observables	4
III. Generalized Hamilton–Jacobi Formalism	10
IV. Classical Model with Global Reparametrization Symmetry	19
V. Physical Interpretation of Extended Momentum	26
VI. Relational Quantization	28
VII. Simple Quantum Model	33
VIII. Conclusion	35
Funding	36
Acknowledgements	36
References	36

I. INTRODUCTION

This paper address the ‘problem of time’ in theories invariant under global time reparametrizations.¹ We provide a new prescription for quantization, leading to a new identification of the quantum observables. Within §II-III, our proposal is motivated by an analysis of symmetry and observables within a generalised Hamiltonian-Jacobi formalism. In §IV, a simple model is used to illustrate how our approach relaxes the usual Dirac prescription [4, 5] for the observables of a totally constrained system (‘perennials’) and allows one to recover the influential partial and complete observables approach [6–11] in a particular limit.² Difficulties such as the non-unitary evolution of the complete observables in terms of certain partial

¹ For a comprehensive overview of the various aspects of the ‘problem of time’ for theories invariant under *local* time reparametrizations, see [1–3]

² For a review of the various notions of observable, that includes discussion of the limitations of the partial and complete observables approach, see [12].

observables are explained as a breakdown of this limit. Identification of our observables (‘mutables’) relies upon a physical distinction between gauge symmetries that exist at the level of histories and states (‘Type 1’), which are formally identified by symmetries generated by linear constraints, and those that exist at the level of histories and not states (‘Type 2’), which are not generated by linear constraints.³ This distinction resolves a tension in the literature concerning the physical interpretation of the partial observables and allows for a richer class of observables in the quantum theory.

One question of central concern is the classical limit of our new quantum formalism: it might be feared that the relaxation of the Hamiltonian constraint equation involved in our proposal would prevent the recovery of the correct behaviour semi-classically. This point is addressed at length in the text below. Within §V, we consider important questions relating to role of conserved quantities in our framework. In particular, we consider the implications of the fact that the quantum operator conjugate to the time parameter within our formalism must correspond to a quantity that is classically conserved. Then, in §VI, we use the well-established connection between the quantum formalism and the Hamilton-Jacobi formalism to argue that the evolution equation we propose for the quantum theory will correspond to the evolution equation for the Hamilton-Jacobi functional established in §III.

Another potential worry regarding our approach is the range of applicability: since we restrict to theories invariant under global time reparametrizations one might be concerned that application of our approach to gravitational models is ruled out in principle. However, as discussed in §VIII: i) there is precedent for the gravitational application of our proposal in terms of the mini-superspace model for uni-modular gravity given by Wald and Unruh [14];⁴ and ii) there is the potential for the application of our proposal to the quantization of gravity when understood in terms of the Shape Dynamics formalism [15–17], which describes the evolution of gravitational degrees of freedom in terms of a global time.

³ As discussed in §II, our approach is related to earlier work of Kuchař [2, 13].

⁴ In the mini-superspace approximation, uni-modular gravity does not suffer from the objection that the uni-modular condition does not seem to be compatible with foliation invariance. Incidentally, away from this approximation, where this objection applies in General Relativity, the application of Relational Quantization to Shape Dynamics is perfectly valid because Shape Dynamics has a notion of global time.

II. SYMMETRY AND OBSERVABLES

Identification of the physical quantities that can be measured – the ‘observables’ – within a constrained Hamiltonian theory generally follows a prescription, due to Dirac [4], whereby observables are defined as functions on the physical phase space that (weakly) commute with the first class constraints. The various quantization techniques applicable to these theories are then required to preserve this notion of observable in the quantum context [5]. However, the use of *Dirac observables* is problematic in totally constrained systems, such as general relativity, within which the Hamiltonian is itself a constraint.⁵ In such theories, the Dirac observables are usually interpreted as unchanging ‘perennials’, leading to an acute problem in recovering a description of physical change.⁶ Several promising lines of argument towards the construction and interpretation of perennials exist in the literature. In particular, the partial and complete observables approach [6–11] provides an algorithm for the construction of Dirac observables in totally constrained systems.⁷

Here we take a different approach to the problem of observables and change in totally constrained systems, but one which contains the Dirac observables as a special case. On our view, one must distinguish between two notions of gauge invariance often taken as equivalent in the literature. The first notion was emphasised in Dirac’s original work [4] and arises when there exists an equivalence class of *states* in the state space of a theory that are identified as *physically indistinguishable*, leading to an under-determination problem in the evolution equations of the theory. A second notion of gauge invariance arises when there is an equivalence class of *histories* in the space of allowable histories of a theory.⁸

A symmetry at the level of states in the state space is not defined primarily in terms of an invariance of the action (although it will always result in one). Rather, such symmetries are understood to result directly from redundancy within the representation of the freely specifiable initial data of the system. As an example, take the case of the ‘hidden’ Weyl invariance of General Relativity that becomes manifest when expressed in terms of Shape

⁵ These difficulties become noticeably worse in the context of chaotic systems [18]. An interesting line of future work would be to evaluate the relative merits of the mutables and complete observables approaches in the context of the classical and quantum issues with chaotic systems.

⁶ Here we should note that it is not universally agreed that the Dirac observables of totally constrained systems must be understood as perennials. See [19] for an alternative evaluation.

⁷ The extent to which Dirac and complete observables need be taken to be synonymous is a further interesting, and to our knowledge unsettled, question.

⁸ Both here and below, ‘history’ refers to a path within the state-space, and is not intended to have any history-theoretic connotations. The notion of a ~~4~~history symmetry is closely connected to Bergmann’s notion of a gauge symmetry. See [12, 20, 21] .

Dynamics [15]. In this case, a parametrization of the gauge-invariant degrees of freedom of the theory can be found in a special choice of foliation where the physical degrees of freedom can be expressed solely in terms of Weyl invariant quantities, even though this Weyl invariance is not an explicit symmetry of the action of general relativity.

A symmetry at the level of histories, on the other hand, is defined directly in terms of an invariance of the action. Any state symmetry implies the existence of a history symmetry as a special case, since a history is just a sequence of states. However, the converse is not true. Even in well-known cases of gauge theories, such as Maxwell theory or Yang–Mills theory, care must be taken to ensure that a state symmetry is equivalent to a history symmetry when the gauge transformations are time-dependent [19, 22–25].

The distinction between gauge symmetries occurring at the level of ‘histories *and* states’ and gauge symmetries occurring at the level of ‘histories *but not* states’ will be of fundamental importance to the arguments of this paper. We will therefore make the following *informal distinction*: a ‘Type 1’ symmetry exists within a theory when there is a state symmetry with a corresponding history symmetry; a ‘Type 2’ symmetry exists when there is history symmetry with *no* corresponding state symmetry. These informal definitions provides an intuitive picture for how to understand our distinctions between Type 1 and Type 2 symmetries. They also allow us to (again informally) define a ‘Type 3’ symmetry as the situation where there exists a symmetry at the level of spacetimes but *no* corresponding state *or* history symmetries.

	Type 1	Type 2	Type 3
Spacetime Symmetry	✓	✓	✓
History Symmetry	✓	✓	×
State Symmetry	✓	×	×

TABLE I: Three Types of Symmetry

These informal definitions can be made more precise as follows.⁹ First, we define the configuration space in the usual sense as the space of all possible instantaneous configurations of the system (i.e., the configurations at a particular time). Using this, we define

⁹ For an exploration of the basis behind these two notions in terms of the distinction between ‘manifest free’ and ‘manifest fixed’ variations of the action see [26].

the phase space as the cotangent bundle to configuration space. We then assume that the canonical analysis of the theory can be well defined and leads to a canonical Hamiltonian system involving first class constraints. These constraints define a Hamiltonian flow on phase space via the symplectic structure. A Type 1 symmetry is formally defined as a symmetry associated with a constraint that has a well defined action on *configuration space*. Such a constraint will be linear and have a Lie group phase space (and configuration space) action that can be identified as a symmetry transformation reflecting a redundancy in the phase space variables.¹⁰ A Type 2 symmetry is defined as a symmetry associated with a constraint that *does not* have a well defined action on configuration space but *does* have a well defined action on phase space. Such a constraint will be non-linear and have a phase space action that *cannot* be identified as a symmetry transformation in the sense of reflecting a redundancy in the phase space variables. Rather, the constraint will be associated with a Lie group action on the space of parametrized configuration space *curves* and it is this action that can be identified as a symmetry transformation of the theory.¹¹

These formal definitions realise the intuitive picture introduced earlier. If a particular constraint induces an action on configuration space, it means that the symmetry can be defined *at an instant* without reference to any further information about the history of the system. This is precisely the defining characteristic of a Type 1 symmetry. If, on the other hand, a particular constraint does not admit a uniquely determined action on configuration space, this means that more information about the history of the system, not contained in the instantaneous configurations, is required to specify the action of the symmetry. Such symmetries we label Type 2. The prototypical example of a Type 1 symmetry is the $U(1)$ invariance of Maxwell theory while, for a Type 2 symmetry, the prototypical example is reparametrization invariance.

One benefit of the distinction between Type 1 and Type 2 relates to the complete and partial observable program mentioned above. This approach involves us identifying both a set of phase space variables – the ‘partial observables’ – and a set of correlations between these observables – the ‘complete observables’. Whilst it is generally agreed that

¹⁰ By definition, such an action must be generated by a sub-algebra of the full constraint algebra. This would rule out application of this classification scheme to situations where the set of linear constraints do not have this property. See [12] for discussion of this issue in the case of supergravity.

¹¹ For canonical general relativity, it will be convenient to additionally distinguish Type 3 symmetries that do not have a uniquely determined action on phase space and are associated with groupoid transformations rather than Lie group transformations. See the end of this section for discussion of this point.

the complete observables are measurable quantities, the status of the partial observables is controversial. Whereas Rovelli [6–8, 27] identifies these functions with physical quantities that can be measured, Thiemann [11] argues that partial observables should be understood as gauge-dependent quantities that are non-measurable. As shall be detailed below, there are good reasons to endorse the Thiemann perspective for partial observables connected to Type 1 symmetries and the Rovelli perspective for partial observables connected to Type 2 symmetries.

A more significant insight gained from the Type 1 and Type 2 distinction relates to the quantization of reparametrization invariant theories with a global notion of time. The Dirac quantization algorithm implements a kinematical prescription to remove the dependence of both the wavefunction and the observables on the parameters that label the ‘gauge orbits’ connecting physically indistinguishable states. Dirac’s prescription is, thus, explicitly motivated by the identification of the instantaneous states along the gauge orbits. Although such an identification is justified in the case of a Type 1 symmetry, it is certainly *not* justified in the Type 2 case. Thus, on our view, the Dirac quantization algorithm is only justified for the Type 1 case. One then requires a new definition of observables and a new approach to quantization for gauge theories with Type 2 symmetries. This is precisely what the *relational quantization* program, defended here and elsewhere [26, 28], aims to provide.¹²

Below we will describe a new implementation of the relational quantization program and show how it correctly implements Dirac’s requirements when, and only when, they are appropriate. Whilst, in previous work, relational quantization was presented in terms of path integral [28] or constraint quantization techniques [26], here we will emphasise the connection to a generalised Hamilton–Jacobi formalism. One particular feature of this new formalism is that it allows us to make a direct connection between the change of the classical characteristic functional with respect to an independent parameter, and the change of the quantum wavefunction with respect to the same parameter. When such a parameter is associated with a Type 2 symmetry, we take both the classical and quantum changes

¹² ‘Relational quantization’, in our terminology, implies an approach towards the quantization of totally constrained systems within which fundamental time ordering structure is retained. We view any approach that jettisons this structure as ‘timeless’ rather than properly speaking ‘relational’. In this respect we are proposing a brand of ‘quantum relationalism’ that is conceptually distinct from others approaches in the literature. See [7, 29–35] for more on other notions of ‘relationalism’ explored in the quantum context and [36] for more on the philosophical background to our approach. See for [37–39] relevant discussions of the difference between Barbour and Rovelli’s articulation of classical relationalism. Further articulation and evaluation of the relationship between the various classical and quantum forms of relationalism is an interesting project.

to be physical. The change that is parameterized is treated as physical even though the independent parameter labelling the change does not correspond to a measurable quantity. Reparametrization invariance in the presence of a global time label is the most important example of a Type 2 symmetry and, in that case, the relevant change parameter is time. Thus, relational quantization leads to a quantum reparametrization invariant formalism with genuine dynamical change. This is in stark contrast to quantum theories resulting from all other canonical quantization techniques. Within the quantum formalism of standard approaches, a Wheeler–DeWitt-type equation imposes the *kinematical* restriction of invariance of the wavefunction with respect the time parameter. Within a relationally quantized theory, a Schrödinger-type equation gives the *dynamical* evolution of the wavefunction with respect the time parameter.

Since change with regard to the time parameter is treated as physical, relational quantization allows for the possibility of such change within the observables of a totally constrained theory. Our approach implements a dynamical notion of observable that is closely related to that defended by Kuchař [2, 13] (see also [40]). Like Kuchař, we disagree with the ‘perennials’ prescription whereby only the timeless Dirac observables of a totally constrained system have the capacity to represent physical quantities. And, like Kuchař, we endorse an enlarged set of observables that need not commute with all the constraints. However, our prescription is distinct from that of Kuchař since we are more restrictive as to the types of constraints that observables may be non-commuting with respect to.

Let us introduce three notions of observable relevant to totally constrained systems as follows: *Perennials* are functions that commute with all the (first class) constraints; *Mutables* are functions that commute with the (first class) constraints associated with Type 1 symmetries but not those associated with Type 2 symmetries; *Kuchař observables* commute with all the (first class) constraints apart from Hamiltonian constraints. When there is a single Hamiltonian constraint, this constraint is associated with a Type 2 symmetry and the Kuchař observables and mutables are identical. When there are multiple Hamiltonian constraints, in theories such as canonical general relativity, things become more complicated and the two notions come apart.

The infinite family of Hamiltonian constraints in canonical general relativity are connected to *local* time diffeomorphism invariance. Crucially, local time diffeomorphism invariance can-

not be unambiguously classified as either a Type 1 or Type 2 symmetry. This is because the set of local time diffeomorphisms include two very different types of transformations that are not clearly distinguished in the formalism. Canonical general relativity involves specification of geometrical information relating to both canonical data on sequences of spatial hypersurfaces and the embeddings of these hypersurfaces into spacetimes. Local time diffeomorphisms are a set of symmetry transformations that includes both transformations that preserve embeddings (reparametrizations) and transformations that do not (refoliations). The embedding preserving transformations can be unambiguously understood as Type 2 symmetries: it makes sense to talk of reparametrizations as symmetries of histories but not states since they reparametrize phase space curves without changing their image. Refoliations, on the other hand, change the image of phase space curves (like a Type 1 symmetry) but do so without admitting a unique state-by-state representation of their action. This is because, in addition to a history, one needs to specify the embedding of the phase space data into spacetime in order to define a refoliation. Without this embedding information, it is impossible to compare two histories on phase space. Another way of viewing this issue is to consider refoliations as normal deformations of three dimensional hypersurfaces embedded within four geometries [41]. In that context, it is clear that they will require a spacetime metric in order to be defined. This metric can be specified either in terms of spacetime geometric data or via integration of spatial data *and* the necessary smearing functions that specify the embedding. Under either viewpoint, we see that, by definition, refoliations cannot be classified as either Type 1 or Type 2 symmetries. This suggests the identification of a new ‘Type 3’ symmetry consisting of transformations that: i) are symmetries of the action; ii) do not preserve embeddings; and iii) have no unambiguous action on phase space.

Since phase space curves map to curves on configuration space, there are obvious difficulties in the quantization of Type 3 symmetries. In particular it is difficult to see how one can consistently define a wavefunction in a configuration basis unless one also has a way to specify the appropriate embedding information quantum mechanically. We will not attempt to tackle this problem here. Rather, given that refoliation invariance is a symmetry of Type 3, we accept that relational quantization, as well as the associated arguments regarding observables and change, cannot be directly applied to gravity when expressed in the full refoliation invariant formalism of canonical general relativity. Fortunately, considerations

regarding the interpretation of scale suggest an alternative formulation for gravity where instantaneous states can be used to construct complete solutions *without* having to specify the embedding of the hypersurfaces into spacetime. Because this new interpretation of gravity — called *Shape Dynamics* [15, 16, 42, 43] — contains a global notion of time, its symmetries can be unambiguously divided into Type 1 and Type 2, and relational quantization can be applied. In this paper, we will lay the necessary groundwork for considering the case of gravity by showing how relational quantization leads to an explicit characterisation of mutables in models with global reparametrization symmetry. In these cases, it is demonstrated that perennials constructed via the partial and complete observables scheme can be recovered as a subset of the mutables constructed within our approach. Thus, the formalism presented here allows us to: i) give an explicit construction of what constitutes an observable in a relationally quantized framework, and ii) make direct contact between our previous proposal and ‘internal time’ methods for understanding change within fully constrained Hamiltonian systems.

III. GENERALIZED HAMILTON–JACOBI FORMALISM

We begin by presenting a generalized Hamilton–Jacobi formalism for constrained classical systems. This approach is closely related to those found in the literature [7, 44, 45]. In particular, for Type 1 symmetries, our formalism reproduces the usual Dirac formalism; however, in the case of Type 2 symmetries, there are several crucial differences. An explicit application to a particle model that features both globally reparametrization-invariant symmetries, which we take to be Type 2, and Type 1 symmetries will be provided in the following section.

Consider a totally constrained Hamiltonian theory on a phase space Γ , coordinatized in some chart by (q, p) , defined by the first class system of r constraints $C_\alpha(q, p) \approx 0$, for $\alpha = 1, \dots, r$. The canonical action reads

$$S = \int [p \cdot \dot{q} - \lambda^\alpha C_\alpha(q, p)] dt. \quad (1)$$

For the moment, we use an abstract index-free notion for phase space variables (q, p) so that the dot product represents an abstract inner product on Γ . Thus, our considerations will

generally apply to the infinite dimensional case. Integration is over the arbitrary parameter t and over-dots represent t -derivatives. The symplectic term requires Γ to be equipped with the symplectic potential $\theta = p \wedge dq$ which defines a symplectic 2-form $\omega = d\theta$ (where exterior derivatives and wedge products are defined on Γ). In addition, we will assume that Γ is equipped with a canonical Poisson structure such that $\{q, p\} = 1$ in some chart and all other Poisson brackets are zero.

Our goal is to find a canonical transformation that parametrizes the flow of the constraints C_α locally on Γ and allows us, in particular, to restrict this flow to the constraint surface defined by $C_\alpha \approx 0$. To do this, consider the modified action:

$$S_e = \int \left[p \cdot \dot{q} + \dot{\phi}^\alpha \cdot \mathcal{E}_\alpha - \lambda^\alpha (\mathcal{E}_\alpha - C_\alpha(q, p)) \right] dt \quad (2)$$

defined on the *extended* phase space $\Gamma(q, p) \rightarrow \Gamma_e(q, p; \phi^\alpha, \mathcal{E}_\alpha)$ coordinatized in some finite patch by the additional variables $(\phi^\alpha, \mathcal{E}_\alpha)$. The symplectic potential on Γ_e is now

$$\theta_e = p \wedge dq + \mathcal{E}_\alpha \wedge d\phi^\alpha. \quad (3)$$

The symplectic structure induces a Poisson structure $\{\phi^\alpha, \mathcal{E}_\beta\} = \delta^\alpha_\beta$ through the inverse of the symplectic two form $\omega_e = d\theta_e$, so that ϕ^α and \mathcal{E}_α are canonically conjugate. Eventually, we will use ϕ^α to locally parametrize a canonical transformation along the orbits of C_α . The momenta \mathcal{E}_α are easily seen to be constants of motion, since their equations of motion imply

$$\dot{\mathcal{E}}_\alpha = \{\mathcal{E}_\alpha, H\} = 0. \quad (4)$$

For the special initial condition $\mathcal{E} = 0$, we see that the extended theory defined by S_e is classically equivalent to the original theory defined by S .

In order for the constraints $\mathcal{E}_\alpha - C_\alpha(q, p) \approx 0$ to form a first class surface on the extended phase space Γ_e , the functions $C(q, p)$ must be Abelian (i.e., $\{C_\alpha, C_\beta\} = 0$). For simplicity, we will assume that the $C(q, p)$'s have already been *Abelianized*, which is always possible locally on Γ due to Darboux's theorem [46, p.175].¹³ For the quantum theory, it is also possible to

¹³ Darboux's theorem states that, locally on Γ , one can always coordinatize the constraint surface $C_\alpha(p, q) = 0$ in terms of "orthogonal" (i.e., commuting) coordinates. When re-expressed in terms of these coordinates, the constraints are, by definition, Abelian.

perform a local Abelianization (see for instance [47]), although, in practice, this procedure can be exceedingly difficult to perform explicitly. For various systems, this Abelianization might not be globally feasible. Fortunately, for the important case of gravity considered in our companion paper, relational quantization is applied to Shape Dynamics whose generator of evolution is Abelian with respect to all local constraints of the theory. Thus, the Abelian restriction is irrelevant in that case. Moreover, a rigorous application of the complete and partial observables program also requires an Abelianization of the constraints [9–11], so our approach is no more restrictive than standard methods.

We are now in a position to define a canonical transformation that parametrizes the flow of the constraints. This can be achieved by requiring that the new coordinates $(P, Q; \Phi^\alpha, E_\alpha)$ have zero flow under the transformed C_α , so that they are analogous to the ‘initial data’ of standard Hamilton–Jacobi theory. For the old coordinates (q, p) , we additionally require that the momenta \mathcal{E}_α are constrained to be equal to the $C_\alpha(q, p)$. The new symplectic potential $\theta'_e = P \wedge dQ + E_\alpha \wedge d\Phi^\alpha$ must differ from the old one by an exact form dS . This ensures that the symplectic 2-form $\omega_e = d\theta_e = d\theta'_e$ is invariant. Implementing the above requirements, we obtain

$$\begin{aligned} dS(q, Q; \phi^\alpha, \Phi_\alpha) &= \theta_e - \theta'_e \\ &= p \wedge dq + \mathcal{E}_\alpha \wedge d\phi^\alpha - P \wedge dQ - E_\alpha \wedge d\Phi^\alpha \\ &= p \wedge dq + C_\alpha(p, q) \wedge d\phi^\alpha - P \wedge dQ, \end{aligned} \tag{5}$$

where in the last line we set $\mathcal{E}_\alpha = C_\alpha(q, p)$ and $E_\alpha = 0$. The reason for setting $E_\alpha = 0$ is because E_α represents the generator of flow on the transformed coordinates (Q, P) , which we want to vanish. This is completely analogous to requiring that the transformed Hamiltonian vanish in standard Hamilton–Jacobi theory.

We see from this that $S(q, Q; \phi^\alpha, \Phi_\alpha)$ is a type-1 generating functional for a canonical transformation taking the lower case coordinates to upper case ones subject to our requirements. Figure 1 shows how to interpret it in terms of a flow along the integral curves of $C_\alpha(q, p)$.

It is convenient to convert S to a type-2 generating functional $F(q, P; \phi^\alpha, \Phi_\alpha)$ in the (q, p)

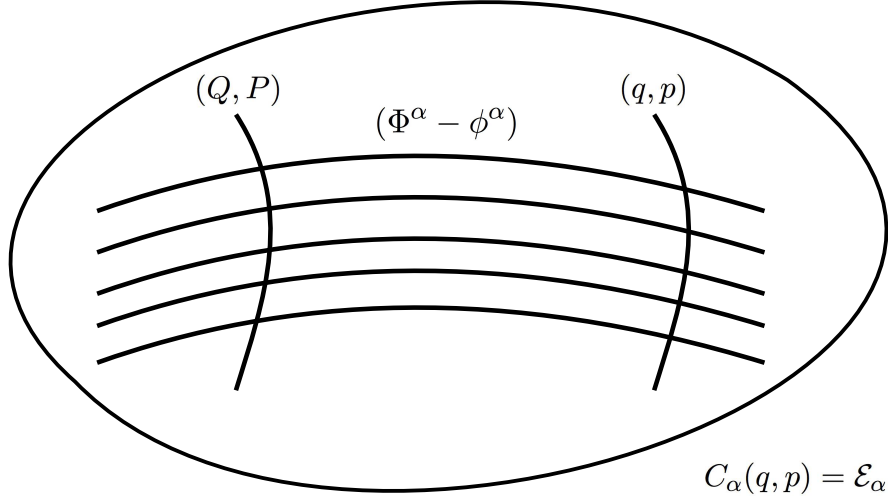


FIG. 1: The canonical transformation gives the general coordinates (q, p) lying along the integral curves of the constraints, $C_\alpha(q, p) = \mathcal{E}_\alpha$, in terms of the ‘initial data’ (Q, P) and the parameters ϕ^α .

coordinates by adding the boundary term $Q \wedge P$. Thus,

$$dF(q, P; \phi^\alpha, \Phi_\alpha) = p \wedge dq + C_\alpha(q, p) \wedge d\phi^\alpha + Q \wedge dP. \quad (6)$$

From the above, we find the equations defining the canonical transformation we are looking for:

$$\frac{\partial F}{\partial q} = p \quad \frac{\partial F}{\partial \phi^\alpha} = C_\alpha(q, p) \quad (7a)$$

$$\frac{\partial F}{\partial P} = Q \quad \frac{\partial F}{\partial \Phi_\alpha} = 0. \quad (7b)$$

The first equation in (7a) is just the definition of the momenta p . Combined with the second equation in (7a), we get the generalized Hamilton–Jacobi relations

$$\frac{\partial F}{\partial \phi^\alpha} = C_\alpha \left(q, \frac{\partial F}{\partial q} \right). \quad (8)$$

The second equation of (7b) simply says that, because (Q, P) are preserved along the flow of the transformed constraints, there is no parameter Φ^α to parametrize their flow. The first

equation of (7b) becomes the Hamilton–Jacobi equation of motion.

We can reduce these equations using the following separation Ansatz for F

$$F(q, P; \phi^\alpha, \mathcal{E}_\alpha) = W(q, P) + \mathcal{E}_\alpha \phi^\alpha, \quad (9)$$

where we have slightly abused notation using the same symbol for the separations constants \mathcal{E}_α as the canonical coordinates \mathcal{E}^α . This abuse is forgivable since they are equal on-shell. Our Ansatz converts the Hamilton–Jacobi relations (8) to the reduced set of equations

$$\mathcal{E}_\alpha = C_\alpha \left(q, \frac{\partial W}{\partial q} \right). \quad (10)$$

Inserting this back into (9), we obtain

$$F(q, P, \phi^\alpha) = W(q, P) + \phi^\alpha C_\alpha \left(q, \frac{\partial W}{\partial q} \right), \quad (11)$$

which gives the generator of the canonical transformation we are looking for parametrized, as advertised, by ϕ^α . Using this, we can compute the final form of the Hamilton–Jacobi equation of motion

$$Q = \frac{\partial F}{\partial P} \Big|_{C_\alpha \left(q, \frac{\partial W}{\partial q} \right) = \mathcal{E}_\alpha}, \quad (12)$$

which should be read as an equation for q in terms of the ‘initial data’ (or reference section), (Q, P) , the constants of motion, \mathcal{E}_α , and the parameters ϕ^α . In this form, it represents an integral of motion for the $q(\phi^\alpha)$ in terms of some initial data.

An alternative formal expression can be given for the integrals of motion of the system in terms of the flow parameters ϕ^α by directly integrating Hamilton’s equations for the extended action in (2). These are the flow equations of the Hamilton vector field of the Hamiltonian. For the extended variables, they are

$$\dot{\phi}^\alpha = \{\phi^\alpha, H\} = \lambda^\alpha \quad \dot{\mathcal{E}}_\alpha = \{\mathcal{E}_\alpha, H\} = 0. \quad (13)$$

Using these, we can rewrite the flow equations for (q, p) as

$$\dot{q} = \{q, H\} = \lambda^\alpha \{q, C_\alpha\} = \dot{\phi}^\alpha \{q, C_\alpha\} \quad (14)$$

$$\dot{p} = \{p, H\} = \lambda^\alpha \{p, C_\alpha\} = \dot{\phi}^\alpha \{p, C_\alpha\} , \quad (15)$$

or

$$\delta q = \delta \phi^\alpha \{q, C_\alpha\} \quad \delta p = \delta \phi^\alpha \{p, C_\alpha\} . \quad (16)$$

We can exponentiate the Lie flow defined by these expressions as:

$$q(\phi^\alpha) = \exp [\phi^\alpha \mathcal{L}_{\mathcal{X}_{C_\alpha}}] \cdot Q = \sum_{n=0}^{n=\infty} \frac{(\phi^\alpha)^n}{n!} \{C_\alpha, Q\}_n \quad (17a)$$

$$p(\phi^\alpha) = \exp [\phi^\alpha \mathcal{L}_{\mathcal{X}_{C_\alpha}}] \cdot P = \sum_{n=0}^{n=\infty} \frac{(\phi^\alpha)^n}{n!} \{C_\alpha, P\}_n . \quad (17b)$$

Where $\{C_\alpha, Q\}_0 := Q$ and $\{C_\alpha, Q\}_{n+1} = \{C_\alpha, \{C_\alpha, Q\}_n\}$. Here Q and P are required to satisfy the initial value constraint $C_\alpha(Q, P) = \mathcal{E}_\alpha$. This form of the integrals of motion is equivalent to solving (12), since both result from the solution of the same variational principle. The form (17) will later allow us to relate our formalism to the complete and partial observables approach. In the quantum formalism, (12) will be analogous to a generalized Schrödinger formalism, while (17) will correspond to a generalized Heisenberg formalism.

The interpretation of the solutions of (12) or (17) depends crucially upon the difference between Type 1 and Type 2 symmetries. Although at the classical level these differences do not have empirical consequences, in the quantum context they lead to different prescriptions for the construction of the physical Hilbert space states and observables (see §VI). For this reason, we will distinguish explicitly between Type 1 symmetries and Type 2 symmetries by using different notation for the associated constraints and conserved charges. For Type 1 symmetries, we will use the subscript β . For Type 2 symmetries we will use the subscript μ . If we assume that s of the r constraints are associated with Type 1 symmetries, we can re-write the constraints C_α , $\alpha = 1, \dots, r$ as (C_β, C_μ) , $\beta = 1, \dots, s$ $\mu = s + 1, \dots, r$. We can then define (ϕ^β, ϕ^μ) and $(\mathcal{E}_\beta, \mathcal{E}_\mu)$ accordingly.

Let us consider Type 1 symmetries first. For these symmetries, the ϕ^β parametrize un-

physical flow. In such circumstances, one must impose upon (12) the additional restriction $\mathcal{E}_\beta = 0$. This will enforce the original constraints $C_\beta = 0$ as the generators of unphysical transformations between states. The condition $\mathcal{E}_\beta = 0$ can be equivalently applied to (17) in terms of a restriction of the initial values of Q and P . There are then two options¹⁴ for the construction of observables in the case of a Type 1 symmetry, both of which lead to Dirac observables:

- i) Retain the parameters ϕ^β and solve (12) (or equivalently (17)) for the integrals of motion, q , as a functions of ϕ^β and some reference section as specified by (Q, P) . This will give different representations of the observables in different gauges labelled by the ϕ^β .
- ii) Solve (17) (or equivalently (12)) to get an explicit expression for the integrals of motion, q , as functions of ϕ^β , Q and P (as per i). Invert s of these expressions for the independent parameters, ϕ^β , in terms of some subset of the integrals of motion for q^β and the (Q, P) . Then, reinsert this result into the remaining equations for the q 's in terms of the q^β and the full set of Q 's and P 's. Observables are then specified as families of functions $F(Q, P, q^\beta)|_{q^\beta=\kappa^\beta}$ where κ^β are a set of real numbers that pick out one member of each family of observables.

The first strategy corresponds to a very general gauge-fixing-type methodology for constructing Dirac observables. The second corresponds to an implementation of the partial and complete observable approach [6, 7, 9–11] with Q and P the partial observables, and $F(Q, P, q^\beta)|_{q^\beta=\kappa^\beta}$ the complete observables. The first strategy has a distinct advantage over the second in that it does not depend upon the invertibility of the integrals motion, which is not generally guaranteed. Nevertheless, when they are well defined, the complete observables will be Dirac observables, and will, for Type 1 symmetries, faithful parameterize all the physical degrees of freedom. The partial observables, on the other hand, can have different values when evaluated at phase space points corresponding to physically identical states. This means that they should not be associated with physical measurements. For Type 1 symmetries, the Thiemann view that partial observables are non-measurable is thus

¹⁴ For explicit discussion of these two options see [7, pp.113-114] and [28, §3.2]. In the context of General Relativity, the second option is closely related to the methodology used by Gerlach [48] to recover the Einstein field equations from the Einstein-Hamilton-Jacobi equation (first written down by Peres [49]).

supported against the Rovelli view that they can be associated with measurements. We will find further support for this conclusion concerning Type 1 symmetries in the context of the model of §IV.¹⁵

Let us next consider the case of Type 2 symmetries. In our view, these symmetries are markedly different to Type 1 symmetries since the associated labelling parameters and conserved charges have physical significance. For the conserved charges, \mathcal{E}_μ , this physical significance is in terms of their role as conserved quantities within the theory that have previously been interpreted as couplings of the theory (see §V for more details). For the labelling parameters, ϕ^μ , the physical significance is more subtle, and is in terms of the *labelling of physically distinct states of the system*. Although, in these senses, the \mathcal{E}_μ and ϕ^μ are physically significant, they do not correspond to independent degrees of freedom. Rather they are independent parameters that parameterize physical differences between states. In order to play this role, \mathcal{E}_μ and ϕ^μ must be retained within the classical and quantum formalisms without an increase in the overall degrees of freedom. Classically at least, this is easily achieved: although we have extended the original phase space of the theory, the presence of the first class constraint $C_\mu(q, p) - \mathcal{E}_\mu = 0$ removes the two extra degrees of freedom, meaning that the total number of degrees of freedom is identical to the original un-extended phase space. Since the conserved charges are physically significant (in the sense described in §V), no additional restrictions $\mathcal{E}_\mu = 0$ should be imposed upon the integral of motion given by (12). We understand the canonical transformation generated by F as reshuffling the degrees of freedom such that ϕ^μ parametrizes a coordinate along the orbits of the constraints C_μ . Thus, the coordinates along the orbits associated with Type 2 symmetries are taken to have physical meaning: they are not ‘gauge’ orbits. We will argue in the next section that this is precisely the case for reparametrization invariant theories, the key exemplar of a theory with Type 2 symmetry. In this case, we have a single constraint C , equivalent to the Hamiltonian of the theory. The Hamiltonian, of course, generates time evolution, so the ϕ associated with the Hamiltonian constraint parametrizes time evolution and the constant of motion \mathcal{E} is the total energy.¹⁶

Observables for theories with Type 2 symmetries are given by the full set of Q s and P s,

¹⁵ See [11, p.78] and [8, 27] for further discussion on this point.

¹⁶ In Jacobi’s action principle, which leads to a Hamiltonian constraint, the energy is considered a coupling of the theory in agreement with our considerations in §V.

with the integrals motion given by solutions of (12) or (17) describing genuine evolution in terms of the independent parameters ϕ^μ . Given these expressions, one has the option to apply the complete and partial observables scheme, which is equivalent to option ii) above. This will serve to eliminate the independent parameters ϕ^μ and give an expression for the relevant Dirac observables. However, such a construction does not give the full set of observables of the theory since there exist measurable quantities which are not described within any single representation of the complete observables. In a system with two configuration space dimensions (i.e., four phase space dimensions) and a single Type 2 symmetry, for example, any construction of the complete observables will represent only half the phase space degrees of freedom, despite there being no redundancy within the theory at the level of states. The partial observables, on the other hand, will parametrize the full phase space, and can in fact be understood as the full set of observables of the theory. Moreover, since there are no ‘gauge orbits’ on phase space, the curves defined by solving (12) or (17) will be unique up to parametrization. For Type 2 symmetries, the Rovelli view that partial observables can be associated with measurements is thus supported against the Thiemann view that partial observables are non-measurable. We will find further support for this conclusion concerning Type 2 symmetries in the context of the model of §IV.

In our scheme, although observables are required to be constant along the flow associated with Type 1 generating constraints (mirroring the usual Dirac requirement), no such requirement is imposed for the Type 2 generating constraints. We will call the class of observables *mutables*. In general, the Dirac observables will be a subset of the mutables.¹⁷ In particular, the Dirac observables can be understood as the subset in which the \mathcal{E}_μ associated with Type 2 constraints is also set to zero and method i) or ii) applied. In the following section, we will provide an explicit application of the generalized Hamilton–Jacobi construction of the mutables and also construct the relevant Dirac observables via the complete observables route. This will then lead naturally into the key physical differences that emerge in the quantum regime — to be discussed in §VI.

¹⁷ An interesting problem for future work would be to establish a precise formal relationship between the mutables and the conception of an observable according to Bergmann [12, 20, 21].

IV. CLASSICAL MODEL WITH GLOBAL REPARAMETRIZATION SYMMETRY

A global time reparametrization transformation induces a mapping between different parametrizations of otherwise identical configuration space curves. Formally, such symmetries can be characterized in terms of the action of the one dimensional Lie group of diffeomorphisms, $\mathcal{D}iff(\mathbb{R})$, upon the space of parametrized configuration space curves, $\mathcal{C}_p \in \{\gamma, t\}$, made up of a pairing of a curve, γ , and a monotonic parameter, t . For the case of a global time parameter t , the action of the reparametrization group, $\mathcal{R}ep$, takes the simple form:

$$\{\gamma, t\} \rightarrow \{\gamma, f(t)\} \quad (18)$$

for $\frac{df(t)}{dt} > 0$. From an abstract mathematical perspective, we can identify $\mathcal{R}ep$ with that of the one dimensional Lie group of diffeomorphisms of the real line \mathbb{R} . By definition a global reparametrization invariant theory is one in which the action is invariant under $\mathcal{R}ep$. Global reparametrization invariance thus requires a total Lagrangian that is homogeneous of order 1 in velocities. This is because, under this requirement, a global reparametrization $t \rightarrow f(t)$ transforms the Lagrangian as $L(q, \dot{q}) \rightarrow L(q, q') \frac{df}{dt}$, where a prime indicates differentiation with respect to f . This cancels the transformation properties of the volume form, $dt \rightarrow \frac{dt}{df} df$, leaving the action invariant:

$$I = \int_{\gamma} L(q, \dot{q}) dt \rightarrow \int_{\gamma} \left(L(q, q') \frac{df}{dt} \right) \left(\frac{dt}{df} df(t) \right) = I, \quad (19)$$

after a relabelling of the dummy integration variable. This requirement on L crucially implies that the momenta will also be invariant under the action of $\mathcal{R}ep$:

$$p = \frac{\partial L(q, \dot{q})}{\partial \dot{q}} \rightarrow \frac{\partial L(q, q') \frac{df(t)}{dt}}{\partial q' \frac{df(t)}{dt}} = p. \quad (20)$$

Thus, instantaneous states, which are defined in terms of points in phase space, are invariant under reparametrization. Nevertheless, it is easy to see that reparametrization invariance *will* always be associated with a canonical constraint. Since the Lagrange density is homogeneous of order 1 in the velocities, Euler's homogeneous function theorem implies that

the Hamiltonian density must vanish; or, rather, that it be proportional to a constraint [4]. Thus, there is a precise formal basis for the existence of (global) Hamiltonian constraints in (global) reparametrization invariant theories. Such constraints have *nothing whatsoever* to do with a transformation between physically indistinguishable instantaneous states. Rather, when considered as acting on phase space points, (global) Hamiltonian constraints have the unambiguous role of producing dynamical transformations between physically distinct states. It is of course true that these transformations are parametrized by arbitrary time labels, however, the image of the phase space curves that are traced out are invariant under reparametrization of the time labels. Thus, we always have an unambiguous sequence of physically distinct phase space points, irrespective of any arbitrariness regarding the time labelling of this sequence.¹⁸

The existence of (global) Hamiltonian constraints is entirely consistent with Type 2 symmetries as defined in both formal and informal senses. Informally speaking, global Hamiltonian constraints are associated with a Type 2 symmetry since reparametrization transformations relate physically indistinguishable histories and there is no corresponding state symmetry relating physically indistinguishable states. Reparametrization symmetry *does not* result from the existence of unphysical directions in the state space. Rather, it results purely from an invariance of the action. Formally speaking, global Hamiltonian constraints are associated with a Type 2 symmetry since they are non-linear constraints that do not have a well defined configuration space action but do have a well defined phase space action.

With these considerations in mind, let us consider a simple theory with both reparametrization Type 2 symmetries and the more typical Type 1 symmetries.¹⁹ The finite dimensional particle model we will study has a total Hamiltonian given by:

$$H(\vec{q}_i, \vec{p}_i) = N\mathcal{H}(\vec{q}_i, \vec{p}_i) + \vec{\lambda} \cdot \vec{\mathcal{P}}(\vec{p}_i), \quad (21)$$

which is defined on the phase space $\Gamma(\vec{q}_i, \vec{p}_i)$, where i ranges over the number of particles

¹⁸ We should note that in the context of canonical general relativity the local Hamiltonian constraints are associated with local time reparametrizations or refoliations which *do* lead to nontrivial transformations of the image of phase space curves. However, this connection can only be established if one considers time-dependent smearings that trace out complete histories. Thus, it is also inappropriate to interpret local Hamiltonian constraints, in and of themselves, as generating transformation between physically indistinguishable instantaneous states.

¹⁹ An alternative example to consider would be the string worldsheet, where spatial diffeomorphisms on the worldsheet would represent a Type 1 symmetry and time reparametrizations on the worldsheet would represent a Type 2 symmetry. Indeed it is well-known²⁰ that these constraints should be considered differently. It would be interesting to compare what is done in this case with what is suggested by relational quantization.

n . In this free ‘Jacobi’ theory, evolution of the particle positions, \vec{q}_i , and momenta, \vec{p}_i , is generated by the Hamiltonian constraint

$$\mathcal{H}(\vec{q}_i, \vec{p}_i) = \sum_i \frac{\vec{p}_i^2}{2m_i} - E \approx 0, \quad (22)$$

which characterizes a Type 2 symmetry, as we’ve just explained. The ‘Gauss-like’ constraint,

$$\vec{\mathcal{P}} = \sum_i \vec{p}_i \approx 0, \quad (23)$$

implements a gauging of the spatial translational invariance of the free particle system. The lapse function N and the vector $\vec{\lambda}$ act as Lagrange multipliers enforcing the vanishing of these constraints. The vanishing of the total linear momentum implies that the position of the centre of mass, $\vec{q}_{\text{cm}} = \frac{1}{\sum_i m_i} \sum_i \vec{q}_i$, is pure gauge. This gauge symmetry is interpreted in terms of a closed system that can only measure inter-particle separations and has no way of physically distinguishing between one value of its centre of mass and another. It is, therefore, characterized unambiguously as a Type 1 symmetry.

This model is particularly useful for understanding canonical general relativity whose constraints also split into an evolution generator (the ADM Hamiltonian constraint) and a generator of an instantaneous symmetry (the ADM Diffeomorphism constraint). It is inspired by more general Barbour–Bertotti models which have been studied extensively [50–53].

Following the general procedure outlined in §III, we define the extended theory

$$S_e = \int dt \left[\sum_i \dot{\vec{q}}_i \cdot \vec{p}_i + \dot{\vec{\sigma}} \cdot \vec{\Upsilon} - \dot{\tau} \mathcal{E} - N(\mathcal{E} - \mathcal{H}) - \vec{\lambda} \cdot (\vec{\Upsilon} - \vec{\mathcal{P}}) \right], \quad (24)$$

where the extended variables $(\tau, \vec{\sigma})$ are arbitrary labels parametrizing the time and centre of mass of the system respectively. The energy, \mathcal{E} , can be thought of as a redefinition of the zero of the total energy of the system $E \rightarrow E + \mathcal{E}$. The other conjugate momentum variable, $\vec{\Upsilon}$, is the total linear moment of the system. We have conventionally added a minus sign to the $\dot{\tau} \mathcal{E}$ term to ensure the usual relations between time and energy. The extended theory is physically indistinguishable from the original in the case of the (τ, \mathcal{E}) extension, and when $\vec{\Upsilon} = 0$ in the case of the $(\vec{\sigma}, \vec{\Upsilon})$ extension.

After making the appropriate identifications, the Ansatz (11) for F , takes the form

$$F(\vec{q}_i, \vec{P}_i; \vec{\sigma}, \tau) = W(\vec{q}_i, \vec{P}_i) + \vec{\mathcal{P}} \left(\vec{q}_i, \frac{\partial W}{\partial \vec{q}_i} \right) \cdot \vec{\sigma} - \mathcal{H} \left(\vec{q}_i, \frac{\partial W}{\partial \vec{q}_i} \right) \tau, \quad (25)$$

where \mathcal{H} and $\vec{\mathcal{P}}$ obey the reduced Hamilton–Jacobi relations

$$\mathcal{E} = \mathcal{H} \left(\vec{q}_i, \frac{\partial W}{\partial \vec{q}_i} \right) \quad \vec{\Upsilon} = \vec{\mathcal{P}} \left(\vec{q}_i, \frac{\partial W}{\partial \vec{q}_i} \right). \quad (26)$$

Using the definitions (22) and (23) and the additional Ansatz for W

$$W(\vec{q}_i, \vec{P}_i) = \sum_i \vec{q}_i \cdot \vec{P}_i, \quad (27)$$

we get

$$F(\vec{q}_i, \vec{P}_i; \vec{\sigma}, \tau) = \sum_i \vec{q}_i \cdot \vec{P}_i - \left(\sum_i \frac{\vec{P}_i^2}{2m_i} - E \right) \tau + \vec{\sigma} \cdot \sum_i \vec{P}_i. \quad (28)$$

The Hamilton–Jacobi equation of motion, (12), for this system is then

$$\frac{\partial F}{\partial \vec{P}_i} = (\vec{q}_i + \vec{\sigma}) - \frac{\vec{P}_i \tau}{m_i} = \vec{Q}_i, \quad (29)$$

which can easily be inverted for \vec{q}_i

$$\vec{q}_i = \left(\vec{Q}_i - \vec{\sigma} \right) + \frac{\vec{P}_i \tau}{m_i}. \quad (30)$$

The \vec{P}_i ’s in this equation must obey the constraints

$$\mathcal{E} = \sum_i \frac{\vec{P}_i^2}{2m_i} - E \quad \vec{\mathcal{P}} = \sum_i \vec{P}_i = 0 \quad (31)$$

in order for the reduced Hamilton–Jacobi equations (26) to be satisfied. Note crucially that, while imposing $\vec{\mathcal{P}} = 0$ is necessary for the classical equivalence of the extended and un-extended theories, no such condition is required on \mathcal{E} , which only contributes a shift of the total energy (more on the interpretation of this in §V). These last equations correspond simply to constraints on the freely specifiable initial momenta.

This theory reproduces precisely our expectations: it gives the usual integral of motion for the free particle plus an extra term that just shifts the origin of each particle system by $\vec{\sigma}$, which is clearly the coordinate along the gauge orbit for this theory. Exploring the various different ways of interpreting the integral of motion (30) gives insight into the nature of observables and how they differ for Type 1 and Type 2 symmetries.

Consider first the Type 1 symmetry associated with the coordinate σ . For this simple symmetry, there is a natural way to parametrize the reduced phase space. If we sum (30) over the mass weighted particle indices and divide by the total mass, $m_{\text{tot}} = \sum_i m_i$, we obtain (upon using the second constraint of (31))

$$\vec{\sigma} = \vec{Q}_{\text{cm}} - \vec{q}_{\text{cm}}, \quad (32)$$

which is just the change in the centre of mass of the system. Reinserting this back into the integral of motion (30) gives

$$\vec{q}_i^{\text{cm}} = \vec{Q}_i^{\text{cm}} + \frac{\vec{P}_i}{m_i} \tau, \quad (33)$$

where the ‘cm’ refers to the use of centre of mass coordinates. Because of the redundancy implied by the translational symmetry, the centre of mass coordinates over-parametrize the reduced phase space. We then have the choice between options i) and ii) detailed above. In this case, i) corresponds to treating (33) as an over-complete set of relational observables. These are easily verified to be Dirac observables for the system. Option ii), on the other hand, leads us to treat the coordinates \vec{q}_i as explicit functions of the initial data (\vec{Q}_i, \vec{P}_i) and the change in centre of mass $\vec{\sigma}$ (and the time, τ). This is equivalent to treating the change in centre of mass as a partial observable that parametrizes the different representations of the coordinates \vec{q}_i in different gauges. Although this can be done for this system, it is very unnatural because it parametrizes things that are not physically measurable — the explicit coordinates, \vec{q}_i — in terms of something else that is not physically measurable — the change in centre of mass, $\vec{\sigma}$. For this example of Type 1 symmetries, the Thiemann perspective, in which partial observables are non-measurable quantities, is more appropriate than the Rovelli perspective, where the partial observables are understood as measurable. Moreover, since it is only the Dirac observables that have operational significance, one could simply follow the more direct definition via option i) and avoid the partial and complete observables

entirely.

Contrast this with what happens when we apply the same options to Type 2 symmetries. Option i) is just equivalent to a standard gauge fixing, which we have the ability to do arbitrarily. For example, the integral of motion in terms of centre of mass coordinates is equivalent to using a gauge fixing $\vec{\sigma} = 0$. The analogous condition $\tau = 0$ indeed gives us a parametrization of the reduced phase space. Then, the ‘integral of motion’ reduces to the trivial statement

$$\vec{q}_i^{\text{cm}} = \vec{Q}_i^{\text{cm}}. \quad (34)$$

The only non-trivial equations in this gauge are, perhaps unsurprisingly, the initial value constraints (31). Clearly this option does not provide us with any dynamical information. We will see in Section 5 that this timelessness is also found when Dirac quantization is applied to this manifestly dynamical model. This should not be surprising since Dirac quantization rests on the assumption of Type 1 symmetries.

The only way to obtain a notion of evolution for this system is to consider (33) as a genuine evolution equation for the system. Rather than taking the timeless route and following option i) towards the construction of Dirac observables for Type 2 symmetries, one follows the alternative mutables route. We can identify the physically relevant observables of the system as those corresponding to the *entire* set of configuration space variables \vec{q}_i^{cm} after removing the centre of mass. These are the mutables for this system since they commute with the ‘Gauss’ constraints \mathcal{P} but not the Hamiltonian constraint \mathcal{H} . These mutables evolve according to (33), tracing out curves labelled by the arbitrary parameter τ , which is of course itself not an observable. Rather, τ is an independent parameter, and, as such, can be specified independently of quantities which are deemed measurable within the theory. The curves defined by (33) are reparametrization invariant even if the equation makes reference to the unphysical labelling parameter. Thus, the mutables, as we have defined them *are* invariant under the relevant Type 2 symmetry.

One might, however, wish to give a ‘parameter free’ expression for the relative variation of the observables. It is in that context that the complete and partial observables program — option ii) — offers a unique strength. Consider our formalism in 1-dimension. We can choose the partial observables to be the centre of mass coordinates, q_i^{cm} , meaning the q_i^{cm} defined via (33), play the role of the ‘flow equations’. A natural choice of clock variables is

the centre of mass coordinate of one of the particles, say q_1^{cm} . We can invert (33) for particle 1 to obtain

$$\tau = \frac{m_1}{P_1} (Q_1^{\text{cm}} - q_1^{\text{cm}}) . \quad (35)$$

The first initial value constraint of (31) gives P_1 in terms of the remaining P_i 's. Using this, we can deparametrize the evolution of the remaining q_a 's (where $a = 2 \dots n$) in terms of q_1

$$q_a^{\text{cm}}(Q_i^{\text{cm}}, P_i, q_1^{\text{cm}}) = Q_a^{\text{cm}} - \frac{m_1}{m_a} P_a \frac{(Q_1^{\text{cm}} - q_1^{\text{cm}})}{\left[2m_1 \left(E + \mathcal{E} - \sum_a \frac{P_i^2}{m_i} \right) \right]^{1/2}} . \quad (36)$$

For any $q_1^{\text{cm}} = \kappa \in \mathbb{R}$ this expression defines a ‘complete observable’, which will also be a Dirac observable. Since these observables are defined as functions of the mutables they constitute a subset of them. In practice, then, one can use the complete and partial observables program to deparametrize the evolution purely in terms of observable quantities. This evolution is, however, fundamentally controlled by (33) and is *always* well-defined, even when a particular deparametrization breaks down. Thus, on our view, even if one wishes to use parameter-free ‘complete observable’ expressions, one is still required to retain the full ‘partial observables’ representation given by (33). This indicates that the the Rovelli perspective, in which partial observables are measurable quantities, is more appropriate than the Thiemann perspective, where the partial observables are understood as non-measurable. Moreover, it coincides with a scheme where the mutables, and not the Dirac observables, are fundamental.

Our example has thus illustrated the following conclusions: 1) standard gauge fixings and parametrizations of the reduced phase space are the appropriate techniques for dealing with Type 1 symmetries; 2) in the case of Type 2 symmetries, the evolution of the system is described by the reduced integral of motion (33) in terms of the arbitrary parameter τ , which is always well-defined; 3) in that context, there are good motivations for applying the complete and partial observables program in order to construct parameter-free complete observable expressions. However, one must retain the parameter dependent ‘partial observable’ expressions in order to fully represent the physics of the system. In §VI, we will argue that such interpretational differences between Type 1 and Type 2 symmetries at the classical level imply a physical difference at the quantum level. This difference can be characterised

in terms of a different behaviour for observables in certain quantum theories, including the model constructed in this section.

V. PHYSICAL INTERPRETATION OF EXTENDED MOMENTUM

Given the extension procedure developed above, it is important to understand the roles played by the extended variables $(\phi^\alpha, \mathcal{E}_\alpha)$. For Type 1 symmetries, this is straightforward since ultimately our treatment of them requires imposing $\mathcal{E}_\beta = 0$. The dynamics of the extended theory is then manifestly equivalent to the original. The ϕ^β become labels that parametrize unphysical flow on the constraint surface.

For Type 2 symmetries, our proposal is to treat the flow parametrized by ϕ^μ as physical. This suggests that we put no additional restrictions on \mathcal{E}_μ since we wish to identify all the original configuration space variables as physical. It is then important to understand the physical significance of the unconstrained constants of motion \mathcal{E}_μ . In general, this will depend on a detailed study of the physics of the system. Here, we will restrict to the case where the only Type 2 symmetry present in the system is that of a reparametrization symmetry. In this case, the relevant Type 2 constraint is a Hamiltonian constraint $C_\mu \rightarrow \mathcal{H}$ and the parameter associated to its flow is the time $\phi^\mu \rightarrow \tau$. We will now see that it is always possible to associate the momentum, \mathcal{E} , with a coupling constant of the theory, which should be interpreted as the total energy.

Consider the decomposition of \mathcal{H} into the general form

$$\mathcal{H} = T(q, p) + \sum_I c^I V_I(q), \quad (37)$$

where $T(q, p)$ is a kinetic term, $V_I(q)$ are different potentials terms, and c^I are coupling constants associated to each $V_I(q)$. The general Hamiltonian for the theory takes the form

$$H = N\mathcal{H} + \lambda^\beta C_\beta. \quad (38)$$

Because the Lagrange multiplier N can take any form, we can define an equivalent Hamiltonian by multiplying the Hamiltonian constraint by an *arbitrary*, nowhere vanishing function, $f(q, p)$, such that $\mathcal{H}' = f(q, p)\mathcal{H}$ provided we redefine N appropriately: $N' = N/f(q, p)$.

By using $f(q, p) = 1/V_i(q)$ for some arbitrarily selected value of i , one can always find an equivalent expression for \mathcal{H} which singles out a particular coupling c^i such that it appears as a constant in the transformed Hamiltonian. The interpretation of this form of the Hamiltonian is that, by changing the lapse, N , you are using a parametrization such that c^i is interpreted as the total energy of the system.

This, it would seem, introduces an ambiguity for the quantization of a reparametrization invariant theory. The reason for this potential ambiguity is that, since any coupling of the theory can be interpreted as an energy for some parametrization, how should one define the vacuum state — which is defined by the state that minimizes the energy — in an unambiguous way? In practice, this is done by privileging a set of inertial observers who see the ‘true’ vacuum of the system. For a harmonic oscillator potential, for example, it is clear which normalisation of the Hamiltonian constraint should be used to define the vacuum of the system: no one could confuse a spring constant for an energy. Thus, in practice, any potential ambiguity should be resolved by the physical properties of the system. In many ways, this situation is analogous of the Unruh effect in quantum field theory, which also privileges the vacuum of a class of inertial observers.

The implications of this for the current discussion follow from the observation that the extension procedure $\mathcal{H} \rightarrow \mathcal{H} - \mathcal{E}$ can be applied equivalently to *any* of the normalizations of the Hamiltonian constraint. The interpretation of such a procedure is to single out one of the coupling constants c_i of the theory and promote it to a constant of motion. Although this should not change the interpretation of the classical theory because it is equivalent to a redefinition of the coupling $c_i \rightarrow c_i + \mathcal{E}$, care must be taken with regard to the implications to the quantum theory. It would be unwise to introduce a procedure that allows, for example, a harmonic oscillator system to be in a superposition of different values of the spring constant. To avoid such a scenario, one must identify the class of inertial observers of the system in question that are used to define the ‘true’ vacuum of the theory. This singles out the unique coupling of the theory that can be associated with the energy of the system. In turn, this fixes the normalization of the Hamiltonian constraint that is to be used to perform the extension procedure required for relational quantization, which we will now describe in detail.

VI. RELATIONAL QUANTIZATION

The most important implication of the foregoing considerations is to be found within the quantum realm. All standard quantization procedures for theories with gauge symmetry — including path integral [54], constraint [55, 56], or reduced phase space [57, 58] approaches — are founded upon a particular identification of the physical degrees of freedom that assumes one must eliminate redundancy within the representation of the physical states. While such an identification is well justified for Type 1 symmetries, it is not for Type 2. The case of reparametrization invariance explicitly illustrates this point since, as discussed above, such a symmetry can only be understood in terms of redundancy at the level of histories, and has no implications for redundancy at the level of states.

The requirements for a quantization technique that faithfully perseveres the physical characteristics of a classical theory with Type 2 symmetries are clear. First, such a technique should take the classical mutables as the basis for the algebra of quantum observables, which are defined as Hermitian operators on a physical Hilbert space. Second, we should define quantum wavefunctions as the rays of this physical Hilbert space that are invariant under change with respect to the independent parameters of whatever Type 1 symmetries might be present in the theory. In other words, we require that Type 1 symmetries be treated equivalently to standard gauge theory techniques. Third, such quantum wavefunctions should evolve according to an evolution equation that reduces in the semi-classical limit to the Hamilton–Jacobi evolution equation (8) with respect to the generator of the Type 2 symmetry. In particular, this should imply that the wavefunction should be able to exist in superpositions of eigenstates of the constant of motion associated with the Type 2 symmetries.

Relational quantization is a technique specifically geared to fulfil these requirements. In previous work, the articulation of relational quantization depended upon an explicit extension procedure combined with standard quantization techniques. Here, we use the generalised Hamilton–Jacobi formalism introduced above to give a more direct presentation. This treatment follows the well-trodden road from the Hamilton–Jacobi characteristic functional to the quantum mechanical wave equation (see for example [59, pp. 99-109]). Essentially, one considers families of hypersurfaces of constant value of the characteristic function as wave-

fronts propagating in configuration space with respect to the independent parameters. The crucial step is then to interpret these wavefronts as surfaces of constant phase of a complex valued wavefunction on configuration space evolving with respect to the independent parameters. In the context of our formalism, this means one takes the functional $F(q, P, \phi^\alpha)$ of (11) as the basis for a complex wavefunction $\Psi(\phi^\alpha)$, defined on a Hilbert space labelled by the eigenvalues of a complete set of the mutables (\hat{q}, \hat{p}) and the independent parameters ϕ^α . For the moment, we will not distinguish between Type 1 and Type 2 symmetries but will discuss the differences later.

As per standard canonical quantization, one takes the symplectic structure of Γ to generalize naturally to the operator algebra $\{q, p\} = \mathbb{1} \rightarrow [\hat{q}, \hat{p}] = i\hbar \hat{\mathbb{1}}$.²⁰ The functional dependence of the wavefunction on the independent variables ϕ^α is given by the equations:

$$C_\alpha(\hat{q}, \hat{p})\Psi = -i\hbar \frac{\partial \Psi}{\partial \phi^\alpha}, \quad (39)$$

which are the direct generalization of (8). We can now posit the separation Ansatz (the analogue of (9))

$$\Psi = \sum_n c_n \exp(-i\hbar \mathcal{E}_\alpha^n \phi^\alpha) \psi_n, \quad (40)$$

where n labels all eigenstates of the operator $C_\alpha(\hat{q}, \hat{p})$

$$C_\alpha(\hat{q}, \hat{p}) \psi_n = \mathcal{E}_\alpha^n \psi_n. \quad (41)$$

These are the analogues of the reduced Hamilton–Jacobi equations (10). As we shall see later, these equations play a similar role within the theory to the time-independent Schrödinger equation.

The appropriate definition of the Hilbert spaces on which these operators act and the construction of the inner product under which observables should be self-adjoint depends upon the physical interpretation of the equations (39).

In the case of a Type 1 symmetry, the change parameterized by the independent parameters is unphysical. This means that (39) should be interpreted as quantum flow equations

²⁰ Strictly speaking a more rigorous algebraic quantization method is required to insure these brackets (as well as the structures below) are well defined. Such formal issues are tangential to our current project and so can reasonably be neglected. See [11, 56, 60–63] for more details.

parametrizing physically indistinguishable quantum states. Resultantly, one should follow the standard Dirac reasoning and impose the following *kinematical* restrictions on the theory: a) physical states of the quantum theory, Ψ_{phys} , should be ϕ^β -independent; and b) the physical Hilbert space should be constructed in such a way that the inner product is invariant under the action of the C_β 's. (Recall that we are again using the index β to label Type 1 symmetries only.)

In the case of a Type 2 symmetry, the change parameterized by the independent parameters is physical. This means that the equations (39) should be treated as *dynamical* equations giving the evolution of the quantum states with respect to the ϕ^μ . (Recall that we are again using the index μ to label Type 2 symmetries only.) This implies no kinematical restriction, and so places no further requirements upon the possible physical states, physical inner product, or observables.

Our analysis of Type 1 symmetries coincides with the standard approach to gauge symmetries. For this case, we can make use of rigorous methodologies within the existing literature to implement conditions a) and b). One such technique is *group averaging*, which is a key aspect of Refined Algebraic Quantization (RAQ) [56]. Essentially, one defines the physical states by averaging Ψ over the group manifold. For simple cases²¹ this averaging takes the form:

$$\Psi_{\text{phys}} = \int d\phi^\beta \Psi \quad (42)$$

$$= \int d\phi^\beta \sum_n c_n \exp(-i\hbar \mathcal{E}_\beta^n \phi_\beta) \psi_n \quad (43)$$

$$= \sum_n c_n \delta(\mathcal{E}_\beta^n) \psi_n \quad (44)$$

$$= c_{0,\beta} \psi_{0,\beta}. \quad (45)$$

The group averaging thus superselects the eigenstate of energy zero, and we recover the results of Dirac's original constraint quantization scheme: i.e. that the physical states should

²¹ For more complex cases, such as where zero is not in the discrete spectrum of the constraint operator, the full machinery of Refined Algebraic Quantization would be needed. The generality of RAQ as a technique for the quantization of theories with Type 1 symmetries extends to all cases where: i) the constraints are self-adjoint operators with an algebra which closes with structure constants; ii) the symmetry group is a finite dimensional locally compact group with respect to a suitable topology. Notably, these conditions are not satisfied for the Hamiltonian constraints of general relativity, where other techniques are required [11].

satisfy

$$C_\beta(\hat{q}, \hat{p})\psi_{\text{phys}} = 0. \quad (46)$$

We can use the group averaging to define a rigging map, η , from the general states Ψ to the physical states. An inner product can then naturally be defined on the physical Hilbert space using this map:

$$\langle \eta(\Psi^1), \eta(\Psi^2) \rangle = \langle \Psi_{\text{phys}}^1, \Psi_{\text{phys}}^2 \rangle = \langle \psi_0^1, \psi_0^2 \rangle. \quad (47)$$

The most important point here is that the construction of the physical Hilbert space involves a quantum analogue of classical phase space reduction. The rigging map is a projection which removes unphysical states, but it also has a further role in ‘quotienting out’ physically identical states. These states lie upon a ‘quantum gauge orbit’ of parametrized by ϕ^β : $\Psi(\phi^\beta) = e^{\phi^\beta \hat{C}_\beta} \Psi(0)$ [64]. This is the direct integration of the quantum flow equation (39) for the Type 1 symmetries. Any two points along this orbit yield the same physical state and so, in group averaging, we are explicitly removing kinematical redundancy at the level of states. This is true both in simple cases of Type 1 symmetries, such as those considered here, or more generally when such symmetries are treated within Refined Algebraic Quantization. We can see this in the latter case explicitly since one variant of the RAQ definition of the physical Hilbert space involves explicitly taking the quotient of the kernel of the rigging map [56]. We thus again see why application of standard gauge quantization techniques to Type 2 symmetries is inappropriate: such techniques assume redundancy at the level of states not histories. Applying a ‘quantum quotienting operation’ to Type 2 symmetries amounts to identifying physically distinct states corresponding to different values of the evolution parameters ϕ^μ . This is to confuse dynamics with kinematics, and change with redundancy.

Once the definition of the physical Hilbert space is given, observables can be constructed by looking for operators that are self-adjoint with respect to this inner product. The kinematic restrictions that we have imposed with regard to Type 1 symmetries but not Type 2 symmetries ensure that such observables will be quantum mutables. Of particular note is the (generalised) Heisenberg picture construction of the mutables. Since the equations (39) are dynamical equations, not kinematic restrictions, we can consider a Heisenberg picture with

the physical change in observables parametrized by the relevant independent parameters ϕ^μ . This leads to the Heisenberg evolution equations for the mutables $\hat{\mathcal{O}}$:

$$i\hbar \frac{d}{d\phi^\mu} \hat{\mathcal{O}} = [\hat{\mathcal{O}}, \hat{C}^\mu], \quad (48)$$

which are the quantum analogues of (16). These exponentiate to the well-known expressions

$$\hat{\mathcal{O}}(\phi^\mu) = e^{-i\hbar\phi^\mu\hat{C}_\mu^\dagger} \hat{\mathcal{O}}(0) e^{i\hbar\phi^\mu\hat{C}_\mu}, \quad (49)$$

which are the analogues of (17).

The above considerations make clear the main difference between relationally quantized theories and Dirac quantized theories. In relational quantization, there is a larger set of observables because these are not required to commute with the Type 2 constraints. These observables evolve according to the Heisenberg equations (48). This implies that we have a different quantum state than that obtained through Dirac quantization. Specifically, in relational quantization the state is allowed to be in a general superposition of eigenstates of the evolution operators C^μ according to (40). While in the classical theory the introduction of the extended variables $(\phi^\mu, \mathcal{E}_\mu)$ did not change the physical predictions of the theory, the same is no longer true at the quantum level: relational quantization leads to physically different predictions for the behaviour of observables as compared with Dirac quantization.

Another way to see this difference is to note that option ii) of §III is no longer available. It is not possible to simply single out one of the mutables, $\hat{\mathcal{O}}$, and invert its complete integral in terms of the evolution parameters ϕ^μ . This is for the simple reason that the observables are now operators, and operator equations can not be solved in terms of parameters. Instead, we can only apply the quantum analogue of option ii) in a dynamical limit of the theory where the Heisenberg state of system is such that some observables $\hat{\mathcal{O}}$ is highly peaked around its expectation value. In this case, one can approximate the evolution of $\hat{\mathcal{O}}$ by taking expectations values of both sides of (48) (formally, this approximation involves ignoring the higher moments of $\langle \hat{\mathcal{O}} \rangle$). Then it is at least possible, although not guaranteed, that this expression can be inverted for ϕ^μ in terms of $\langle \hat{\mathcal{O}} \rangle$. Thus, only in such a limit is it possible to recover something like the quantum analogue of the complete and partial observables

framework.

In this way, not only do we recover a quantum version of the complete and partial observables program, but our approach provides a more general framework for understanding what happens when this program breaks down. From the point of view of relational quantization, non-unitary evolution of the complete observables in terms of a particular partial observable occurs when the semi-classical approximation required to treat that partial observable as a clock breaks down. In our approach, the evolution is always unitary and well-defined provided the physical inner product can be constructed in such a way that the evolution operators \hat{C}_μ are self-adjoint (this is necessary requirement to being able to apply our procedure consistently).

VII. SIMPLE QUANTUM MODEL

Application of relational quantization to the model of §IV is as follows. The quantum version of the generalized Hamilton–Jacobi relations (8) reads

$$-i\hbar \frac{\partial \Psi(\tau, \vec{\sigma})}{\partial \tau} = \left(\sum_i \frac{\hat{p}_i^2}{2m_i} - E \right) \Psi(\tau, \vec{\sigma}) \quad i\hbar \frac{\partial \Psi(\tau, \vec{\sigma})}{\partial \vec{\sigma}} = \left(\sum_i \hat{p}_i \right) \Psi(\tau, \vec{\sigma}), \quad (50)$$

for the wavefunction $\Psi(\tau, \vec{\sigma})$, which represents a state in the auxiliary Hilbert space and is dependent upon both the independent parameters τ and $\vec{\sigma}$.

The Ansatz (25) generalizes to

$$\Psi(\tau, \vec{\sigma}) = \int d\mathcal{E} d^3\vec{\Upsilon} c(\mathcal{E}, \vec{\Upsilon}) \exp \left\{ i\hbar \left(\mathcal{E}\tau + \vec{\Upsilon} \cdot \vec{\sigma} \right) \right\} \psi_{\mathcal{E}, \vec{\Upsilon}}. \quad (51)$$

This leads to the reduced equations

$$\left(\sum_i \frac{\hat{p}_i^2}{2m_i} - E \right) \psi_{\mathcal{E}, \vec{\Upsilon}} = \mathcal{E} \psi_{\mathcal{E}, \vec{\Upsilon}} \quad \left(\sum_i \hat{p}_i \right) \psi_{\mathcal{E}, \vec{\Upsilon}} = \vec{\Upsilon} \psi_{\mathcal{E}, \vec{\Upsilon}}, \quad (52)$$

which are the analogues of (26). The $\psi_{\mathcal{E}, \vec{\Upsilon}}$ are thus energy and total momentum eigenfunctions. These eigenfunctions and their eigenvalues can be constructed in analogy to the

Ansatz (27)

$$\psi_{\mathcal{E}, \vec{\Upsilon}} = \exp \left(-i\hbar \sum_i \vec{P}_i(\mathcal{E}, \vec{\Upsilon}) \cdot \vec{q}_i \right). \quad (53)$$

Here, we have used a configuration basis where

$$\hat{q}_i = \vec{q}_i \quad \hat{p}_i = -i\hbar \frac{\partial}{\partial \vec{q}_i} \quad (54)$$

and the $P_i(\mathcal{E}, \vec{\Upsilon})$ are integration constants depending on $(\mathcal{E}, \vec{\Upsilon})$. Their dependence on $(\mathcal{E}, \vec{\Upsilon})$ is obtained implicitly by the fact that they must obey the relations

$$\mathcal{E} + E = \sum_i \frac{\vec{P}_i^2}{2m_i} \quad \vec{\Upsilon} = \sum_i \vec{P}_i, \quad (55)$$

which are obtained by plugging the Ansatz (53) into the reduced equations (52). Note that the states (53) are all non-normalizable, as is usually the case for the free particle, so that the usual care (i.e., some regularization procedure) must be taken to define an inner product on states.

For the Hamiltonian constraint, no group averaging is required because change parameterized by τ is interpreted as being physical. However, for the total momentum constraint, the transformations parameterized by the $\vec{\sigma}$ are manifestly unphysical. They are simply ‘Leibniz shifts’ expressed as redefinitions of the origin of the centre of mass. Thus, group averaging is required in this case. For this system, this involves constructing the physical states

$$\begin{aligned} \Psi_{\text{phys}}(\tau) &= \int d^3\vec{\sigma} \Psi(\tau, \vec{\sigma}) \\ &= \int d\mathcal{E} c(\mathcal{E}, 0) \exp \left\{ i\hbar \left(\mathcal{E}\tau - \sum_i \vec{P}_i \cdot \vec{q}_i \right) \right\} \end{aligned} \quad (56)$$

where, in the last line, the integration over both $\vec{\sigma}$ and $\vec{\Upsilon}$ simply sets $\vec{\Upsilon} = 0$ (note that the $P_i(\mathcal{E}, \vec{\Upsilon})$ still must satisfy (55)). The inner product on physical states can now similarly be defined through the relevant rigging map.

The mutables are all operators constructed from the centre of mass operators $\hat{q}_i^{\text{cm}} = \hat{q}_i - \frac{1}{M_{\text{tot}}} \sum_i \frac{\hat{q}_i}{m_i}$ (where M_{tot} is the total mass) and the conjugate operators $\hat{p}_i^{\text{cm}} \equiv \hat{p}_i - \frac{1}{n} \sum_i \hat{p}_i$.

Their evolution in terms of τ is given by the Heisenberg evolution equations

$$-i\hbar \frac{d}{d\tau} \hat{\mathcal{O}}(\hat{q}_i^{\text{cm}}, \hat{p}_i^{\text{cm}}, \tau) = [\hat{\mathcal{O}}, \hat{\mathcal{H}}] \quad (57)$$

where

$$\hat{\mathcal{H}} = \sum_i \frac{\left(\hat{p}_i^{\text{cm}}\right)^2}{2m_i} - E. \quad (58)$$

These clearly exceed the number of Dirac observables, which must also commute with the Hamiltonian constraint $\hat{\mathcal{H}}$ and, thus, can have no evolution with respect to it. Furthermore, the wavefunction, Ψ , of this theory can be formed from an arbitrary superposition of energy eigenstates according to (56). Note that these states *can* be made normalizable unlike the pure states one obtains from Dirac quantization.

VIII. CONCLUSION

In this paper, we first isolated the important physical distinction between Type 1 gauge symmetries, which exist at the level of histories and states, and Type 2 gauge symmetries, which exist at the level of histories and not states. This distinction was characterised explicitly using a generalized Hamilton–Jacobi formalism within which a non-standard prescription for the observables of classical totally constrained systems was developed. These ideas were then used to motivate a ‘relational quantisation’ procedure which is different from Dirac quantization. In particular, relational quantization of totally constrained systems was shown to lead to a formalism that allows for superpositions of energy eigenstates and includes an enlarged set of quantum observables: the ‘mutables’. Explicit construction of these observables and characterisation of their dynamical evolution can be achieved for simple totally constrained models. An immediate and tractable extension is to apply relational quantization to mini-superspace models. There, one has good reason to expect that the mutables prescription for observables combined with the extra freedom to take superpositions of energy eigenstates will allow for phenomena not permitted within the Dirac framework. Indeed, for mini-superspace models, relational quantization is formally equivalent to the uni-modular model briefly presented by Wald and Unruh at the end of section III of [14]. Developing this model in more detail is an exciting prospect for future work.

Before concluding we should again emphasise an important qualification regarding the limits of our proposal. Our approach is based upon the assumption that, when they exist, state symmetries can be unambiguously identified at the level of phase space. This is explicitly not the case for theories with refoliation symmetry such as General Relativity. Without further articulation, refoliation symmetries cannot be fit into either the Type 1 or Type 2 categories. The issue with refoliation symmetry does not, however, entirely bar the application of relational quantization to gravity. This is because, as noted in [26, 28], the technique can be applicable to the Shape Dynamics formulation of gravity [15]. In this formulation, gravitation is no longer foliation invariant. The symmetries can be clearly distinguished as Type 1 spatial diffeomorphism and Weyl symmetries, and Type 2 reparamterization symmetry. Explicit details of this approach and further consideration of the issues of quantization and refoliation invariance will also be addressed in future work.

FUNDING

K.T. would like to thank the Alexander von Humboldt Foundation and the Munich Center for Mathematical Philosophy (Ludwig-Maximilians-Universität München) for support. S.G. would like to acknowledge support from the Netherlands Organisation for Scientific Research (NWO) (Project No. 620.01.784) and Radboud University.

ACKNOWLEDGEMENTS

We are appreciative to Oliver Pooley, Henrique Gomes and Lee Smolin for valuable discussion, and to Edward Anderson and Martin Bojowald for detailed critical comment on an earlier draft.

-
- [1] C. Isham, “Canonical quantum gravity and the problem of time,” *Arxiv preprint gr-qc* (1992) . <http://arxiv.org/abs/grqc/9210011>.
 - [2] K. Kuchař, “The problem of time in canonical quantization of relativistic systems,” in *Conceptual Problems of Quantum Gravity*, A. Ashtekar and J. Stachel, eds., p. 141. Boston

- University Press, 1991.
- [3] E. Anderson, “Problem of time in quantum gravity,” *Annalen der Physik* **524** no. 12, (2012) 757–786.
 - [4] P. A. M. Dirac, *Lectures on Quantum Mechanics*. Dover Publications, Yeshiva University, New York, 1964.
 - [5] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*. Princeton University Press, 1992.
 - [6] C. Rovelli, “Partial observables,” *Phys. Rev. D* **65** (2002) 124013.
 - [7] C. Rovelli, *Quantum Gravity*. Cambridge University Press, Cambridge, 2004.
 - [8] C. Rovelli, “Comment on ”are the spectra of geometrical operators in loop quantum gravity really discrete?” by b. dittrich and t. thiemann,” 2007. <http://arxiv.org/abs/0708.2481>.
 - [9] B. Dittrich, “Partial and complete observables for canonical general relativity,” *Classical and Quantum Gravity* **23** (2006) 6155.
 - [10] B. Dittrich, “Partial and complete observables for hamiltonian constrained systems,” *General Relativity and Gravitation* **39** (2007) 1891.
 - [11] T. Thiemann, *Modern canonical quantum general relativity*. Cambridge University Press, 2007.
 - [12] E. Anderson *et al.*, “Beables/observables in classical and quantum gravity,” *SIGMA* **10** no. 092, (2014) 092.
 - [13] K.uchař, “Time and interpretations of quantum gravity,” in *4th Canadian Conference on General Relativity* World Scientific Singapore, 1992.
 - [14] W. G. Unruh and R. M. Wald, “Time and the Interpretation of Canonical Quantum Gravity,” *Phys.Rev.* **D40** (1989) 2598.
 - [15] H. Gomes, S. Gryb, and T. Koslowski, “Einstein gravity as a 3D conformally invariant theory,” *Class. Quant. Grav.* **28** (2011) 045005, [arXiv:1010.2481](https://arxiv.org/abs/1010.2481) [gr-qc].
 - [16] E. Anderson, J. Barbour, B. Z. Foster, B. Kelleher, and N. O’Murchadha, “The physical gravitational degrees of freedom,” *Class. Quant. Grav.* **22** (2005) 1795–1802, [arXiv:gr-qc/0407104](https://arxiv.org/abs/gr-qc/0407104).
 - [17] S. Gryb and F. Mercati, “2+1 gravity on the conformal sphere,” *Phys. Rev. D* **87** (2013) 064006, [arXiv:1209.4858](https://arxiv.org/abs/1209.4858) [gr-qc].

- [18] B. Dittrich, P. A. Hoehn, T. A. Koslowski, and M. I. Nelson, “Chaos, dirac observables and constraint quantization,” [1508.01947](https://arxiv.org/abs/1508.01947). <http://arxiv.org/abs/1508.01947>.
- [19] J. Pons, D. Salisbury, and K. A. Sundermeyer, “Observables in classical canonical gravity: folklore demystified,” *Journal of Physics A: Mathematical and General* **222** no. 12018, (2010) .
- [20] P. G. Bergmann, “Observables in General Relativity,” *Rev. Mod. Phys.* **33** (1961) 510–514.
- [21] P. G. Bergmann and A. Komar, “The coordinate group symmetries of general relativity,” *International Journal of Theoretical Physics* **5** no. 1, (1972) 15–28.
- [22] J. Pons, D. Salisbury, and L. Shepley, “Gauge transformations in the lagrangian and hamiltonian formalisms of generally covariant theories,” *Phys. Rev. D* **55** no. 2, (1997) 658–668.
- [23] J. Pons and D. Sailsbury, “Issue of time in generally covariant theories and the komar-bergmann approach to observables in general relativity,” *Phys. Rev. D* **71** (2005) 124012.
- [24] J. B. Pitts, “A first class constraint generates not a gauge transformation, but a bad physical change: the case of electromagnetism,” *Annals of Physics* **351** (2014) 382–406.
- [25] J. B. Pitts, “Change in hamiltonian general relativity from the lack of a time-like killing vector field,” *Studies In History and Philosophy of Modern Physics (forthcoming)* (2014) . <http://philsci-archive.pitt.edu/10094/>.
- [26] S. Gryb and K. Thébault, “Symmetry and evolution in quantum gravity,” *Foundations of Physics* **44** no. 3, (2014) 305–348.
- [27] C. Rovelli, “Why gauge?,” *Foundations of Physics* **44** no. 1, (2014) 91–104.
- [28] S. Gryb and K. Thébault, “The role of time in relational quantum theories,” *Foundations of Physics* (2011) 1–29.
- [29] D. Page and W. Wootters, “Evolution without evolution: Dynamics described by stationary observables,” *Physical Review D* **27** no. 12, (1983) 2885–2892.
- [30] C. Rovelli, “Is there incompatibility between the ways time is treated in general relativity and in standard quantum mechanics,” in *Conceptual Problems of Quantum Gravity*, A. Ashtekar and J. Stachel, eds. Birkhauser, 1991.

- [31] L. Smolin, “Space and time in the quantum universe.,” in *Conceptual problems of quantum gravity*, A. Ashtekar and J. Stachel, eds., vol. 1, pp. 228–291. 1991.
- [32] M. Bojowald, P. A. Höhn, and A. Tsobanjan, “An effective approach to the problem of time,” *Classical and Quantum Gravity* **28** no. 3, (2011) 035006.
<http://stacks.iop.org/0264-9381/28/i=3/a=035006>.
- [33] E. Anderson, “The Problem of Time and Quantum Cosmology in the Relational Particle Mechanics Arena,” [arXiv:1111.1472](https://arxiv.org/abs/1111.1472) [gr-qc].
- [34] E. Anderson, “Relationalism,” 05, 2012. <http://arxiv.org/abs/1205.1256>.
- [35] J. Barbour, M. Lostaglio, and F. Mercati, “Scale anomaly as the origin of time,” *General Relativity and Gravitation* **45** no. 5, (2013) 911–938.
- [36] S. Gryb and K. P. Y. Thébault, “Time remains,” *British Journal for the Philosophy of Science* (**forthcoming**) (2015) .
- [37] K. P. Thébault, “Symplectic reduction and the problem of time in nonrelativistic mechanics,” *The British Journal for the Philosophy of Science* **63** no. 4, (2012) 789–824.
- [38] K. P. Thébault, “Three denials of time in the interpretation of canonical gravity,” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **43** no. 4, (2012) 277–294.
- [39] E. Anderson, “Machian time is to be abstracted from what change?,” *arXiv preprint arXiv:1209.1266* (2012) .
- [40] J. Barbour and B. Z. Foster, “Constraints and gauge transformations: Dirac’s theorem is not always valid,” *ArXiv e-prints* (Aug., 2008) , [arXiv:0808.1223](https://arxiv.org/abs/0808.1223) [gr-qc].
- [41] C. Teitelboim, “How commutators of constraints reflect the spacetime structure,” *Annals of Physics* **79** no. 2, (1973) 542 – 557.
<http://www.sciencedirect.com/science/article/pii/0003491673900961>.
- [42] J. Barbour and N. O’Murchadha, “Conformal Superspace: the configuration space of general relativity,” [arXiv:1009.3559](https://arxiv.org/abs/1009.3559) [gr-qc].
- [43] H. Gomes and T. Koslowski, “The Link between General Relativity and Shape Dynamics,” *Class.Quant.Grav.* **29** (2012) 075009, [arXiv:1101.5974](https://arxiv.org/abs/1101.5974) [gr-qc].
- [44] C. Rovelli, “A note on the foundation of relativistic mechanics. i: Relativistic observables and relativistic states,” *arXiv preprint gr-qc/0111037* (2001) .

- [45] J. Souriau, *Structure of dynamical systems: a symplectic view of physics*. Birkhauser, 1997.
- [46] R. Abraham and J. E. Marsden, *Foundations of Mechanics*. Benjamin/Cummings, 1978.
- [47] S. Gogilidze, A. Khvedelidze, and V. Pervushin, “On Abelianization of first class constraints,” *J.Math.Phys.* **37** (1996) 1760–1771, [arXiv:hep-th/9504153 \[hep-th\]](#).
- [48] U. H. Gerlach, “Derivation of the ten einstein field equations from the semiclassical approximation to quantum geometrodynamics,” *Phys. Rev.* **177** (1969) 1929–1941.
- [49] A. Peres, “On cauchy’s problem in general relativity-ii,” *Il Nuovo Cimento Series 10* **26** no. 1, (1962) 53–62.
- [50] J. B. Barbour and B. Bertotti, “Mach’s Principle and the Structure of Dynamical Theories,” *Proc. R. Soc. A* **382** no. 1783, (1982) 295–306.
- [51] E. Anderson, “Foundations of Relational Particle Dynamics,” *Class. Quant. Grav.* **25** (2008) 025003, [arXiv:0706.3934 \[gr-qc\]](#).
- [52] E. Anderson, “Relational particle models. I: Reconciliation with standard classical and quantum theory,” *Class. Quant. Grav.* **23** (2006) 2469–2490, [arXiv:gr-qc/0511068](#).
- [53] E. Anderson, “Relational particle models. II: Use as toy models for quantum geometrodynamics,” *Class. Quant. Grav.* **23** (2006) 2491–2518, [arXiv:gr-qc/0511069](#).
- [54] L. D. Faddeev, “The feynman integral for singular lagrangians,” *Theoretical and Mathematical Physics* **1** no. 1, (1969) 1–13.
- [55] P. A. M. Dirac, *Lectures on quantum mechanics*. Dover Publications, 1964.
- [56] D. Giulini and D. Marolf, “On the generality of refined algebraic quantization,” *Classical and Quantum Gravity* **16** (1999) 2479–2488.
- [57] J. Śniatycki and A. Weinstein, “Reduction and quantization for singular momentum mappings,” *Letters in mathematical physics* **7** no. 2, (1983) 155–161.
- [58] T. Thiemann, “Reduced phase space quantization and dirac observables,” *Classical and Quantum Gravity* **23** no. 4, (2006) 1163.
- [59] H. Rund, *The Hamilton-Jacobi theory in the calculus of variations: its role in mathematics and physics*. Van Nostrand London, 1966.
- [60] C. Isham and A. Kakas, “A group theoretical approach to the canonical quantisation of gravity. i. construction of the canonical group,” *Classical and Quantum Gravity* **1** no. 6, (1984) 621.

- [61] C. Isham and A. Kakas, “A group theoretical approach to the canonical quantisation of gravity. ii. unitary representations of the canonical group,” *Classical and Quantum Gravity* **1** no. 6, (1984) 633.
- [62] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*. Univ. Pr., Princeton, USA, 1992. 520 p.
- [63] D. Giulini and D. Marolf, “A uniqueness theorem for constraint quantization,” *Classical and Quantum Gravity* **16** (1999) 2489–2505.
- [64] A. Corichi, “On the geometry of quantum constrained systems,” *Classical and Quantum Gravity* **25** no. 13, (2008) 135013. <http://stacks.iop.org/0264-9381/25/i=13/a=135013>.